Facial expression and selective attention
Patrik Vuilleumier

Recent findings demonstrate that faces with an emotional expression tend to attract attention more than neutral faces, especially when having some threat-related value (anger or fear). These findings suggest that discrimination of emotional cues in faces can at least partly be extracted at preattentive or unconscious stages of processing, and then serve to enhance awareness and behavioural responses toward emotionally relevant stimuli. Functional neuroimaging results have begun to delineate brain regions whose response to threat-related expressions is independent of voluntary attention (e.g. amygdala and orbitofrontal cortex), and other regions whose response occurs only with attention (e.g. superior temporal and anterior cingulate cortex). Moreover, visual responses in the fusiform cortex are enhanced for emotional faces, consistent with their greater perceptual saliency. Recent data from event-related evoked potentials and neurophysiology also suggest that rapid processing of emotional information may not only occur in parallel to, but promote a more detailed perceptual analysis of, sensory inputs and thus bias competition for attention toward the representation of emotionally salient stimuli.

Introduction

Other people’s faces constitute major landmarks in our environment. In a single glance, a unique amount of information can be extracted almost effortlessly from an individual face, including familiarity or identity, age, gender, attractiveness, ethnic origins, direction of gaze, oral speech cues, and emotional state. Facial expressions of emotion are particularly important to guide social and motivational activities. Darwin [1] hypothesized that such expressions have specifically evolved to provide powerful signals for rapid nonverbal communication. Recent research has corroborated these intuitions in many respects, by showing not only that dedicated brain systems allow an efficient and relatively automatic detection of emotional information in faces, but that this may effectively act to control attention and behaviour.

On the one hand, given many simultaneous competing events and a limited processing capacity, our brain needs mechanisms of attention to select those stimuli that are most relevant for current goals or immediate survival. Selective attention is necessary for accurate discrimination and conscious perception. Without attention, even suprathreshold stimuli will escape awareness [2]. A large body of work in neuroscience indicates that attention mechanisms involve large-scale networks in the parietal and frontal cortices which can modulate sensory processing in early modality-specific areas, to selectively enhance the representation of relevant stimuli and suppress irrelevant information [3*]. On the other hand, adaptive behaviour requires that some degree of stimulus processing may still take place independently from attention and awareness in order to guide attention to salient events, even when these occur outside the current focus of attention [4]. From this perspective, a rapid detection of stimuli with emotional significance would have obvious advantages, especially when signalling potential threats. Emotional information might therefore be expected to have a privileged status in the capability of receiving some ‘preattentive’ or unconscious analysis, and be more likely to attract attention than other stimuli.

Several lines of evidence suggest that faces and emotional expressions are endowed with such a special significance for our visual system. Faces may capture attention more readily than other meaningful objects present in the scene [2,5*,6,7]. Furthermore, some emotional facial expressions can be discriminated in a fairly automatic manner: without attention, without
intention, and without conscious knowledge; and such emotional discrimination can in turn enhance attention toward a face, in much the same automatic and involuntary manner. This review will summarize recent findings in support of this idea, which have not only tremendously accumulated over the last years, through a variety of different approaches, but also begun to provide remarkable insights into the underlying neural substrates of such effects.

**Behavioural observations in healthy human subjects**

Differential effects of emotional faces on attention have been observed in a number of typical psychophysical paradigms. Four recent studies using visual search tasks [8–10,11] showed that schematic faces with a negative (i.e. angry) expression were found more quickly than similar faces with a positive (i.e. happy) expression (Fig. 1). In visual search, the time to detect a given target usually increases as a linear function of the number of distractors, indicating serial attentive inspection of every stimulus in the display. However, if target recognition can occur without attention at earlier parallel processing stages, then attention is automatically drawn to the target, ‘pop-out’, and detection times are largely independent of the number of distractors. A faster detection of negative faces therefore suggests that their potential threatening value can be extracted prior to selective spatial attention and can heighten their saliency. All these studies employed schematic faces instead of real photographs with the aim of minimizing potential confounds due to physical differences in image properties. Although ecologically impoverished, such stimuli are known to effectively convey subjective emotional meaning and produce reliable face-specific neural responses in functional neuroimaging [12] and electrical potentials [13], similar to veridical photographs. Moreover, in two studies [8*,9*], an advantage for detecting angry faces was not obtained when stimuli were inverted or when only some features were presented in isolation, suggesting that the effect did not result from just low-level cues (e.g. downward mouth curves) but rather require some holistic processing of the whole face configuration [14*]. In the third study [10*], this advantage of angry faces was also found to be best driven by a specific configuration of downward frowning eyebrows and upwardly curved mouth traits, as opposed to any other combination. However, in this study, the effect persisted even with face inversion, suggesting either some contribution of low-level features, or a special robustness of this prototypical affect representation, surviving despite an inverted configuration [15,16]. Altogether, these findings provide novel support to a seminal but controversial study by Hansen and Hansen [17] who reported that an angry face could pop-out among a set of happy distractors (i.e. producing identical detection times in the presence of six versus nine items), whereas happy faces had no such advantage among angry distractors. However, this so-called ‘face-in-the-
crowd’ effect had subsequently been challenged by several studies which showed mixed findings, no such pop-out, or evidence that the result could be accounted for by visual artefacts in the photographs used [18]. Faster detection of happy faces among neutral distractors has been found only in a few search studies [15] where low-level confound could not be ruled out.

However, the new studies [8,10,11] have also qualified earlier work by showing that the enhanced detection of angry faces did not afford true pop-out effects as in the traditional sense (i.e. flat inspection time slopes irrespective of the number of distractors), but rather yield a more efficient search with shallower inspection slopes and quicker spotting of the target location, relative to faces with neutral or positive expressions (Fig. 1).

A further consistent replication across these visual search studies [8,10,15,17] is that ‘absent’ responses are often significantly slower on displays with only angry faces than on displays with only neutral or only happy faces (Fig. 1). This would also be compatible with attention being captured to a greater extent and ‘dwelling’ longer on any individual face with a negative expression.

Other investigators used a dot probe detection task to examine spatial orienting of attention toward emotional stimuli. Subjects have to detect a peripheral target in the right or left visual field, preceded by a brief presentation of faces at each location, one with a neutral expression and one with an emotional expression. Although these faces are irrelevant to the task, detection of the target tends to be faster when it appears on the same side as a preceding angry face, rather than on the opposite side [19,20]. Such effects again reveal an automatic processing of emotional expression and a spontaneous bias to orient attention toward the location of negative faces. Previous research has suggested that such a bias may be greater in the left visual field; enhanced with presentation of very brief masked faces (14 ms) that are below the threshold for conscious perception; and increased by high or medium levels of anxiety, even in a non-clinical range [19], but more so in general anxiety disorder [21]. More recent results indicate that these effects can occur without overt eye movements [20], but nonetheless correlate with a tendency to make rapid gaze shifts toward the angry face location even in subjects who do not show any significant effects on their manual detection reaction times [22]. This suggests that faces with angry expressions may elicit a rapid and involuntary initial orienting of attention, irrespective of anxiety, while a greater maintenance and longer dwelling of attention on them might be necessary to produce the faster responses in manual detection times associated with anxiety. Indeed, another recent study [23] found that angry faces did not speed up the detection of targets appearing on the same side, but instead slowed detection of targets on the opposite side [24], suggesting a greater advantage in holding rather than attracting attention. These effects were present irrespective of the time interval between face and target (100 or 250 ms) in high-anxious subjects, but only at short intervals in low-anxious subjects. However, since attention was cued by unilateral faces (unlike the bilateral pair in previous studies), it is possible that any advantage of angry faces in initial shifting of attention might have been obscured by much larger non-specific spatial cueing effects also produced by neutral and happy faces. This would suggest that a rapid preferential orienting of attention to threat signals may become especially apparent in situations with multiple simultaneous stimuli, which are more demanding on attention selection [3,25].

Faces with happy or sad expressions do not produce similar spatial orienting in such paradigms, including in people with low anxiety or depression [20,22]. This suggests that these effects may not reflect non-specific emotional or arousal facilitation, mood-congruent biases, or general property of any negative facial expression.

Attentional capture by threatening faces is also exemplified by interference effects in a Stroop-like paradigm, in which subjects must name the colour of photographs depicting faces with either neutral or angry expressions. Colour naming latencies are slower for angry faces, even when the emotional expression is very briefly presented (30 ms) and masked by a neutral face [26,27], suggesting unconscious processing of the irrelevant emotional expression and involuntary diversion of selective attention away from colour. The extent of interference has been reported to be unrelated to anxiety but proportional to high anger personality traits [27], as well as correlated with individual hormonal factors such as cortisol and testosterone levels [26,28]. A similar interference by emotion was observed when subjects performed same–different identity judgements on pairs of faces with fearful or neutral expressions [29,30].

### Neuropsychological observations in brain-damaged patients

Damage to right parietal lobe regions in humans may cause hemispatial neglect, characterized by a deficit in directing attention toward the contralesional (i.e. left) space and a loss of awareness for stimuli on that side. The patients often present with perceptual extinction on bilateral stimulation, whereby they may fail to detect a left-side stimulus when shown with a simultaneous stimulus on the right side, although they can still detect the same left-side stimulus shown alone [30]. Such extinction is thought to result from abnormal spatial bias in attention due to impaired influences on sensory
processing caused by unilateral parietal damage, and thus preventing awareness of contralesional stimuli in the presence of competing ipsilesional inputs [31,32]. However, contralesional extinction is not only less severe for faces than for other meaningful objects [6], but also less severe for faces with angry or happy emotional expressions than for neutral expressions [33\*] (Fig. 2). These findings imply that (1) facial features and emotional expressions are still discriminated from other stimuli despite inattention to contralesional events in these patients; (2) such discrimination can modulate the spatial distribution of attention, enhancing the detection of stimuli with a more salient emotional value; and (3) these effects may persist after unilateral damage to attention mechanisms mediated by the parietal cortex.

Blindsight is a distinct disorder associated with damage to the primary visual cortex and hemianopia, in which unconscious processing of emotional information has been reported for faces shown in the blind visual field [34,35\*].

**Specialized neural systems for processing emotional expression in faces**

Classic work by Ekman and Oster [36] has suggested the existence of six basic facial expressions of emotion (happiness, anger, fear, sadness, disgust, surprise), universally recognized across different cultures. There is also evidence that human babies are efficient at orienting to faces and discriminating between distinct facial emotions [37] very early on, suggesting partly innate hard-wired mechanisms dedicated to face and expression processing in the brain. Functional neuroimaging in healthy subjects and neuropsychological studies of selective deficits in the recognition of facial emotions after brain damage have recently allowed a better delineation of neural systems that may be involved in representing separate categories of emotion [38**..39**].

The best identified circuit from a neurophysiological perspective in both animals and humans is certainly the fear system, in which the amygdala in the anterior medial temporal lobe plays a central role [40]. The amygdala is reliably activated by faces with fearful expressions [41,42*,43–46], or by angry faces with aversive value acquired through fear-conditioning (e.g. pairing with a loud noise burst) [24**,47], even when the emotional face is masked and not consciously perceived [47,48*,49]. Unconscious amygdala activation by fearful faces was also recently reported in a patient with blindsight [50\*]. Such unconscious fear processing might rely on direct pathways from the thalamus, providing crude but fast inputs for threat-related cues bypassing slower analysis in the visual cortex [40].

Angry or happy faces rarely activate the amygdala [51*,52–54] but may engage the anterior cingulate, the orbitofrontal, and the inferior prefrontal cortices [39*. Many of these areas have reciprocal connections with widespread cortical regions involved in the control of attention [24**,25,55].

**Functional neuroimaging of emotion and attention interactions**

In many neuroimaging studies [29**,42*,44,46,47], neural responses to fearful expressions were obtained despite the fact that emotion had no bearing with the subjects’ task (e.g. judging gender), consistent with the idea that emotion might be automatically processed without selective attention or intention. A few recent studies [51*,56*,57] directly compared the effects of other facial expressions across different task conditions requiring.
either explicit processing of emotion (i.e. judging expressions) or incidental processing (i.e. judging gender or identity). A consistent finding emerging from these studies is that explicit attention to emotional expression activates the superior temporal gyrus and inferior prefrontal cortex irrespective of expression type, consistent with a role of these regions in higher-order analysis of changeable aspects of faces [58]. On the other hand, Critchley et al. [51] found that bilateral posterior amygdala/hippocampal regions, the extrastriate visual cortex, the retrosplenial cortex, and the pulvinar nucleus of the thalamus responded to emotional more than neutral faces irrespective of the task, with posterior amygdala/hippocampal activation being even greater during incidental rather than explicit conditions. However, the reported amygdala effects were anatomically imprecise, and emotional expressions were shown in blocks of mixed angry and happy faces. Gorno-Tempini et al. [56] found an emotional response to happy and disgusted expressions only during explicit (but not incidental) processing (unlike earlier studies with fearful faces [29,47]). Altogether, these data suggest that emotional processing may be dependent on selective attention to expressions for happy or disgusted faces, though perhaps not for threat-related signals associated with fearful and angry faces.

Another recent imaging study [29] systematically compared the effects of fearful expression produced by faces that were presented either inside or outside the focus of attention. Subjects had to concentrate their attention on specified peripheral locations in order to perform a difficult matching task for pairs of stimuli briefly shown there, in the presence of task-irrelevant stimuli at other locations (Fig. 3a,b). Faces or houses could unpredictably appear at the relevant or irrelevant locations, while the faces had either fearful or neutral expressions. Face-specific areas in the fusiform extrastriate visual cortex [12] showed increased activity when faces appeared at the relevant/attended locations (Fig. 3c). However, an amygdala activation by fearful compared with neutral expressions occurred regardless of whether the faces appeared at the relevant/attended or irrelevant/ignored locations (Fig. 3d). This provides a direct demonstration that amygdala responses to threat-related stimuli may be independent of attention (perhaps partly mediated by direct subcortical inputs). Increased responses to fear with attention occurred in anterior cingulate and anterior temporal cortex.

In addition, activation of the fusiform cortex was generally higher for fearful than for neutral faces, even when the faces were at irrelevant/ignored locations (Fig. 4). Such findings would be consistent with feedback projections from the amygdala onto extrastriate cortex [59], which might enhance the processing of stimuli with emotional value and operate separately from the modulation of spatial attention. Similarly, a previous study [60] reported increased coupling of amygdala and fusiform activity with fearful versus happy faces. In monkeys, face-selective neurons also exhibit enhanced responses to emotional compared with neutral faces, only 50 ms after their first discriminatory activation [61**]. Such modulation from amygdala (and perhaps other limbic areas) on visual cortex might provide a neural substrate for the enhanced detection of threatening faces, and their more efficient capture of attention among competing stimuli, even when initially appearing outside the focus of attention.

Enhanced visual processing though direct amygdala influences on the visual cortex, independent of spatial attention, might also account for better detection of emotional faces in patients with neglect and extinction, still operating despite their parietal lobe lesion. Indeed, a recent functional magnetic resonance imaging study found that fearful faces still activated the amygdala and orbitofrontal cortex both when consciously seen and when extinguished [62]. Fusiform activity increased with conscious perception compared with extinction, but also increased for fearful faces over neutral faces, irrespective of awareness, a pattern similar to the findings in normal subjects [29**].

**Rapid emotional processing revealed by evoked potentials**

Event-related potentials offer unique tools to assess how fast emotional expressions are processed. There has been a recent increase in such studies. Relatively late components over posterior regions may discriminate emotional from neutral faces around 250 ms after onset, and between fearful and other expressions at around 550 ms [63,64]. However, some of these late components may occur only with selective attention to expression [63,65]. Remarkably, the face-specific N170 wave [13] does not seem to be affected by emotion [66]. There is, however, some indirect evidence for earlier components associated with emotional processing, starting in more anterior and midline brain regions between 100 and 200 ms after onset, even during incidental exposure [64,67,68]. Moreover, short-latency (120–160 ms) responses selective for fearful faces, against happy faces, have been recorded by direct intracerebral electrodes in the medial ventralprefrontal cortex of an awake human subject [69**]. Angry faces can also produce cross-modal enhancement for very early potentials evoked by auditory stimuli [70], consistent with a fast and automatic interaction of emotional information with attentive cognitive processes [71].

These findings together with earlier magnetoencephalographic results [72,73] seem compatible with rapid emo-
tional inputs into anterior limbic regions of the brain (such as the orbitofrontal and cingulate cortices, perhaps via amygdala), followed by reentrant modulatory feedback onto posterior sensory cortices, boosting detection of and attention to emotionally relevant stimuli. The lateral prefrontal cortex might also be implicated [74,75]. Note that the deep and nuclear disposition of the amygdala would make it unlikely to be recorded by surface potentials.

**Implications for psychiatry**

Abnormal processing of facial expressions has been increasingly reported in many major psychiatric diseases.
Enhanced attention biases toward threat cues (e.g., angry faces) have been repeatedly found in non-clinical anxiety and may play an important role in the maintenance of general anxiety disorder [20*,22,23,76]. In keeping with this, functional neuroimaging studies have shown increased amygdala activation in anxiety, phobia, and posttraumatic disorders [77,78*], including in response to fearful faces that were masked and not consciously perceived [48*]. This might lead to enhanced vigilance and prolonged engagement of attentive processes on stimuli perceived as threatening.

By contrast, depression is not associated with increased orienting of attention toward angry or sad faces, nor away from happy faces in dot-detection probe tasks [22]. However, a recent report [11] suggested that in visual search, depressed subjects may be slower in detecting happy faces compared with normal participants, though still faster in detecting angry among neutral faces. An increased amygdala response to fearful faces has been reported in depression, resolving with treatment [49,79]. Since depression is commonly associated with anxiety, it is possible that amygdala responses to negative emotional expressions might be enhanced but that additional dysfunction might reduce a subsequent efficient engagement of attention and alertness.

Conclusions
There is a remarkable convergence of behavioural and neurophysiological evidence suggesting that our brain is equipped with mechanisms enhancing the detection of and reaction to emotional facial information, producing involuntary effects on attention and other behavioural responses [80*,81**]. Such modulations appear most effective for the processing of negative, threat-related expressions (fear and anger), but further research is needed to clarify the effect of other expressions and the role of specific facial cues. These modulations illustrate a dynamic interplay between perceptual processes in sensory cortices, amygdala, and ventral and medial prefrontal areas, beyond traditional views of purely modular and feed-forward processing. Similar interactions may account for enhanced attention to other non-facial stimuli with threat-related emotional value [24**,82–84,85**], although facial expressions may represent unique ‘biologically prepared’ cues that have particular importance for attention.

References and recommended reading
Papers of particular interest, published within the annual period of review, have been highlighted as:
• of special interest
** of outstanding interest

   This is an excellent review of current thinking on the mechanisms of attention at the neurophysiological level. It presents evidence suggesting that top-down signals from a distributed network of areas in frontal and parietal cortex provide biasing influences to resolve the competition among multiple inputs for representation within the visual cortex.


12 Sagiv N, Bentin S. Structural encoding of human and schematic faces irrespective of display exposure (300 versus 800 ms), and a good review of previous work.


Facial emotion and attention

Vuilleumier

43 Shows that fearful faces differentially activate amygdala and medial frontal cortex only when also eliciting arousal skin conductance responses, but hippocampal and lateral frontal activity is in the absence of skin conductance responses. Emphasizes interactions of distinct areas in functional networks.
50 Patients with posttraumatic stress disorder showed a general increase in amygdala responses to faces with angry versus happy expressions that were masked by a neutral face and not consciously seen. This amygdala hyperresponsivity was correlated with severity of posttraumatic stress disorder.
52 This paper reports differential amygdala activation to faces with fearful expressions or fear-conditioned faces with angry expressions presented in the blind visual field of patient GY with unilateral occipital damage and blindsight.
54 This is the first whole-brain imaging study to examine differences in responses evoked by emotional faces when the task demands explicit attentive processing of expression, versus incidental processing while attention is engaged on gender. Mixed blocks of angry and happy faces did not allow determining whether incidental effects were common to the two expressions or not.
60 Emotional neural responses to happy and disgusted faces found only when the task demands explicit attentive processing of expression, but not when attention is directed to gender. This contrasts with other studies showing emotional responses to fearful faces even without attention to expression.
66 Neurons in the superior temporal sulcus of monkeys responded to faces with emotional expression or known identity by increasing their rate of firing and amount of transmitted information after a short initial non-differential delay. These increases in neuronal feedback from remote brain areas and show that the same neurons code different facial information at different latencies.
69 This shows that fearful and happy faces modulate late evoked-potential responses over posterior brain regions compared with other expressions, but only when the task requires explicit attention to expression.
71 Shows a relatively early enhancement of posterior visual potentials to fearful and happy faces relative to neutral faces, with independent component analysis suggesting a coupling with more anterior potentials at the same latency.
75 Compared faces classified as “liked” or “disliked” by each observer, instead of classical emotions. Differential responses to disliked faces occurred very early, around 80–116 ms in right hemisphere and slightly later around 104–160 ms in left hemisphere, with more anterior initial distribution for disliked faces.
78 Selective responses to aversive visual stimuli (faces or scenes) occurred with a very short-latency after onset (120–160 ms), shorter than intracranial potentials associated with face or object recognition [see [66]]. This result suggests that some information about stimulus valence is extracted in anterior limbic areas before complete analysis in visual cortex.
80 The study shows that N100 potentials to the same auditory fragments of angry voices are enhanced when presented after a emotionally congruent (angry) versus incongruent (sad) expressions. It suggests that the combination of emotional information across the two modalities is a mandatory and rapid process, already influencing auditory processing 110 ms after stimulus onset.

This is a useful review of extant data on fear processing in humans and animals, emphasizing the relevance for anxiety disorders.


Shows that people inspect longer the eye regions in faces exhibiting a fake ‘social’ smiling expression (‘Duchenne smile’), i.e. with a similar lower but different upper facial muscle pattern. This is a new demonstration that attention is automatically drawn to emotionally relevant cues in facial stimuli.

Brief masked presentations of happy or angry faces elicited differential mimicry reactions in observers’ facial muscles, corresponding to those involved in generating happy (i.e. zygomatic) and angry (i.e. eyebrows) expressions, respectively, despite unawareness of the masked faces. Such findings provide evidence that emotional information may not only be quickly and unconsciously extracted, but also quickly and unconsciously influence motor behaviour.


Healthy subjects missed words with aversive content less often than neutral words in an attentional blink paradigm, in which successive targets must be identified. Patients with left or right amygdala resections did not show such a benefit for aversive words, consistent with a critical role of the amygdala in enhancing perception of affectively valenced stimuli.